

Année de l'AAP : 2007

Projet N° 07004 Achevé

Titre du projet : Reconnaissance automatisée des espèces adventices des rizières de Camargue par les techniques de reconnaissance de contenus visuels
--

Unité responsable du projet : AMAP (botanique et bioinformatique de l'Architecture des Plantes) (CIRAD, CNRS, INRA, IRD, UMII)

Porteur de projet : Daniel Barthélémy (barthelemy(a)cirad.fr)

Pays associés au projet : Philippines, Inde, Colombie, Afrique de l'Ouest, Madagascar

Unités de recherche du réseau scientifique d'Agropolis Fondation associés : Innovation

Sous-axes thématiques : BIP-1: Génétique et génomique, amélioration des plantes, écophysiologie, BIP-2: Maladies et ravageurs des plantes, protection intégrée des cultures, écologie des populations, DSTI-1: Innovations agro-environnementales, agro-écosystèmes, gestion des ressources

Objectifs:

Les difficultés rencontrées pour l'identification taxonomique sont reconnues comme l'un des freins majeurs à l'application de la Convention sur la Diversité Biologique. Depuis que ce frein a été mis en évidence par la Conférence des Parties, de nombreux efforts ont été réalisés pour diffuser l'information taxonomique ainsi que des outils d'identification plus rapides, pertinents et efficaces.

De nombreuses bases de données taxonomiques textuelles et photographiques ont ainsi été générées et sont diffusées au travers des NTIC (Nouvelles Technologies de l'Information et de la Communication). Cependant leur exploitation reste difficile en raison de leur très grande taille et de leur complexité.

Ce projet a pour objectif de permettre l'identification automatique des adventices des rizières de Camargue par le développement d'algorithmes adaptés et l'emploi d'un système de reconnaissance d'images, couplé à une base de données taxonomiques botaniques. Nous effectuerons nos travaux sur une flore limitée sur laquelle nous disposons déjà d'une grande expertise et de plusieurs milliers d'images. Ces recherches seront concentrées sur les espèces présentes dans l'ouvrage « *Les plantes des rizières de Camargue* », coédité par le Cirad, le Parc naturel régional de Camargue, et le Centre français du riz.

Ceci doit permettre d'analyser la faisabilité de l'approche sur un cas concret, et de développer des méthodes génériques d'identification automatique d'espèces végétales à travers l'emploi d'un système informatique de reconnaissance de contenu visuel. Les résultats obtenus fonderont les bases de projets ultérieurs plus ambitieux portant sur l'identification automatisée des adventices des cultures tropicales, des plantes envahissantes ou des espèces végétales constituant divers écosystèmes naturels notamment tropicaux.

Une base de données photographiques portant sur des illustrations standardisées de 178 espèces végétales adventices des rizières de Camargue sera constituée. Cette base de données, ainsi que celle déjà existante sur cette flore, seront exploitées par les outils de reconnaissance d'images pour optimiser les performances de ces derniers. Les outils actuels seront testés et améliorés à l'aide de séries de tests menés conjointement entre informaticiens, botanistes et malherbologues. Les résultats obtenus feront l'objet de publications conjointes entre les différents partenaires, et une restitution des bases de données, et des outils développés sera diffusée sur les sites Internet des différents organismes participants.

Actions menées et résultats obtenus :

Une phase préliminaire de ce projet a consisté à mettre en place un protocole d'acquisition d'images de plantes, adapté à l'évaluation du moteur de recherche de contenus visuels Ikona. La réalisation de ce protocole de prise de vues, ainsi que de celui du traitement informatique des résultats a été effectué sur une base d'images d'Orchidées du Laos.

Une seconde phase du projet a ensuite consisté à adapter celui-ci à la prise de vues d'adventices des rizières de Camargue, au cours de missions régulières de terrain de juin à octobre 2008. L'ensemble des images réalisées ont ensuite été annotées, puis transmises à l'équipe-projet Imedia qui les a indexées et mises à disposition à travers l'outil de reconnaissance de contenus visuels Ikona. Des tests en fonction de la pertinence d'Ikona, du type de prises de vues, et des espèces photographiées ont alors été menés.

Ces tests ont montré une pertinence élevée du moteur de recherche Ikona sur ces bases d'images. Les bases d'images les plus standardisées (telles que les scans de feuilles), montrent les résultats les plus pertinents, tandis que ce sont celles de photos réalisées sur le terrain qui montrent les résultats les plus faibles, notamment en raison de la plus grande surface et hétérogénéité du fond des images acquises. Ces résultats s'expliquent par l'emploi des descripteurs globaux d'images qui prennent en compte la totalité des informations présentes dans les images. Des tests conduits avec des descripteurs locaux (tels que les points d'intérêts) ont montré dans certains cas de bons résultats, grâce au fait qu'ils permettent de s'affranchir de ces contraintes de fonds, cependant il n'existe jusqu'à présent pas d'outils mathématiques adaptés à leur évaluation pour l'identification taxinomique.

Tous les résultats du projet ont été mis en ligne sous forme de tableaux, et graphiques, permettant de visualiser la pertinence des différents descripteurs, sur les différents types d'images. Deux interfaces distinctes, permettent d'utiliser en ligne le moteur de recherche Ikona sur les bases d'images produites dans le cadre de ce projet. L'interface spécifique à la recherche sur la base d'images d'adventices permet de plus d'accéder pour chaque image aux pages descriptives des espèces présentes sur le site des adventices des rizières de Camargue.

Article :

Le Bourgeois T., Bonnet P., Edelin C., Grard P., Proserpi J., Théveny F., Barthélémy D., 2008. L'identification des adventices assistée par ordinateur avec le système IDAO. *Innovations Agronomiques*, 3 : 167-175

Perspectives :

Les résultats fournis dans le cadre de ce projet, permettent d'orienter les recherches futures pour l'adaptation des outils de reconnaissance visuelle pour l'identification taxinomique, très largement exploités dans le cadre du projet PI@ntNet accepté pour financement. Des travaux sont ainsi actuellement en cours, pour évaluer un descripteur de forme spécifique sur les images de scans de feuilles produites, et en comparer les résultats avec ceux obtenus avec les descripteurs visuels génériques évalués. Les requêtes effectuées avec des descripteurs locaux d'images, semblent offrir des perspectives intéressantes pour l'extension de cette approche, ils doivent pour cela être améliorés pour (i) en permettre leur exploitation et évaluation à plus large échelle, (ii) exploiter de manière simultanée des descriptions de formes et de couleurs, (iii) permettre la description de formes en fonction de leur agencement géométrique. Nos travaux ont également révélé l'intérêt de pouvoir effectuer des requêtes combinées de textes et d'images, ainsi que de pouvoir faire des requêtes multi-classes. Le système de recherche visuel sera en effet d'autant plus pertinent qu'un utilisateur affinera sa recherche sur un type d'image et/ou un groupe taxinomique. La recherche multi-classes (à savoir la possibilité de rechercher une espèce par plusieurs types d'images comme des scans de feuilles et des photos de fleurs) fera l'objet de recherches plus approfondies dans le cadre du projet PI@ntNet. Les techniques d'apprentissages n'ont pas pu être évaluées dans le cadre de ces travaux, elles semblent pouvoir nettement améliorer les réponses du système, c'est pourquoi le bouclage de pertinence avec des points d'intérêts fait l'objet d'un sujet de recherche en thèse, financé par l'Inra et l'Inria.

Ces résultats encourageants sur des espèces d'adventices nécessitent des recherches complémentaires, mises en oeuvre dans le cadre de PI@ntNet. Elles doivent permettre la réalisation d'un outil d'identification multi-approches (visuelle, graphique, textuelle) des adventices des rizières irriguées du monde, ceci à travers le projet PI@ntWeed. Les approches préconisées

durant ces recherches se veulent les plus génériques possibles, c'est pourquoi les résultats de ces travaux serviront également de base pour l'extension de cette approche pour l'identification des plantes et/ou maladies des plantes, par plusieurs types d'images et de caractères d'identification.

Financement total par Agropolis Fondation : 15 547 € (frais de mission, CDD de 3 mois pour un ingénieur, achat de matériel photographique)

Catégorie(s) de soutien : soutiens divers (soutien à des projets exploratoires, risqués et innovants, « *proof of concept studies* », « *new frontier research* »)

Durée du projet: 7 janvier 2008 - 7 janvier 2009

Mots clés: taxonomie – rizière – adventices – photographie – base de données